Notes on the QR algorithm

Robert van de Geijn

Draft November 14, 2000

Assumptions:

$$AS = \Lambda S$$

where

- $$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

  with $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_{n-1}| > |\lambda_n|$. (Note: we are treating only a special case. Generalization is a bit messy, so we skip it.)

- S has linearly independent columns. In other words, A has simple eigenvalues.
- $$S = \begin{pmatrix} s_1 & s_2 & \cdots & s_n \end{pmatrix}$$
- $||s_i||_2 = 1$, $i = 1, \ldots, n$

Exercise 1 *Show that* $As_i = \lambda_i s_i$

1 Power Method

Given initial vector x iterate as follows:

- $x_0 = x$
- do $i = 0, \ldots$ until convergence
  - $\hat{x}_{i+1} = Ax_i$
  - $x_{i+1} = \hat{x}_{i+1}/\lambda_1$

- enddo

Exercise 2 *Show $x = \alpha_1 s_1 + \cdots + \alpha_n s_n$. (Hint: $Sy = x$ has a solution. How do $\alpha_i s$ relate to y?)*

Assume for the rest of the argument that $\alpha_1 \neq 0$. In other words, x has a component in the direction of s_1.

Exercise 3 *Show*

$$Ax = \alpha_1 \lambda_1 s_1 + \cdots + \alpha_n \lambda_n s_n$$

and

$$A^k x = \alpha_1 \lambda_1^k s_1 + \cdots + \alpha_n \lambda_n^k s_n$$

Exercise 4 *Show*

$$x_k = \alpha_1 s_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k s_n \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1}\right)^k s_n$$

Exercise 5 *Show*

$$||x_k - \alpha_1 s_1||_2 \leq (|\alpha_2| + \cdots + |\alpha_n|) \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k$$

and thus

$$||x_{k+1} - \alpha_1 s_1||_2 \leq \left(\frac{|\lambda_2|}{|\lambda_1|}\right) (|\alpha_2| + \cdots + |\alpha_n|) \left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k$$

Notice that thus the bound on the difference between x_k and $\alpha_1 s_1$ (an eigenvector associated with λ_1) is decreased at each step by a factor $\frac{|\lambda_2|}{|\lambda_1|}$.

Practical Power Method:

- $x_0 = x$
- do $i = 0, \ldots$ until convergence
  - $\hat{x}_{i+1} = Ax_i$
  - $x_{i+1} = \hat{x}_{i+1} / ||\hat{x}_{i+1}||_2$
- enddo

2 Inverse Power Method

Given initial vector x iterate as follows:

- $x_0 = x$
- do $i = 0, \ldots$ until convergence
  - $\hat{x}_{i+1} = A^{-1} x_i$

- $x_{i+1} = \lambda_n \hat{x}_{i+1}$

- enddo

Assume for the rest of the argument that $\alpha_n \neq 0$. In other words, x has a component in the direction of s_n.

Exercise 6 *Show*

$$A^{-1}x = \alpha_1 \frac{1}{\lambda_1} s_1 + \cdots + \alpha_n \frac{1}{\lambda_n} s_n$$

and

$$A^{-k}x = \alpha_1 \frac{1}{\lambda_1^k} s_1 + \cdots + \alpha_n \frac{1}{\lambda_n^k} s_n$$

Exercise 7 *Show*

$$x_k = \alpha_1 \left(\frac{\lambda_n}{\lambda_1}\right)^k s_1 + \alpha_2 \left(\frac{\lambda_n}{\lambda_2}\right)^k s_n \cdots + \alpha_{n-1} \left(\frac{\lambda_n}{\lambda_{n-1}}\right)^k s_{n-1} + \alpha_n s_n$$

Exercise 8 *Show*

$$||x_k - \alpha_n s_n||_2 \leq (|\alpha_1| + \cdots + |\alpha_{n-1}|) \left(\frac{|\lambda_n|}{|\lambda_{n-1}|}\right)^k$$

and thus

$$||x_{k+1} - \alpha_n s_n||_2 \leq \left(\frac{|\lambda_n|}{|\lambda_{n-1}|}\right) (|\alpha_1| + \cdots + |\alpha_{n-1}|) \left(\frac{|\lambda_n|}{|\lambda_{n-1}|}\right)^k$$

Notice that thus the bound on the difference between x_k and $\alpha_n s_n$ (an eigenvector associated with λ_n) is decreased at each step by a factor $\frac{|\lambda_n|}{|\lambda_{n-1}|}$.

3 Shifted Inverse Power Method

Let the shift $\sigma \neq \lambda_n$ and assume $|\lambda_1 - \sigma| \geq |\lambda_2 - \sigma| \geq \cdots \geq |\lambda_{n-1} - \sigma| > |\lambda_n - \sigma|$

Given initial vector x iterate as follows:

- $x_0 = x$

- do $i = 0, \ldots$ until convergence

  - $\hat{x}_{i+1} = (A - \sigma I)^{-1} x_i$
  - $x_{i+1} = (\lambda_n - \sigma)\hat{x}_{i+1}$

- enddo

Assume for the rest of the argument that $\alpha_n \neq 0$. In other words, x has a component in the direction of s_n.

Exercise 9 *Show* $(A - \sigma I)s_i = (\lambda_i - \sigma)s_i$

Exercise 10 *Show*

$$(A^{-1} - \sigma I)x = \alpha_1 \frac{1}{\lambda_1 - \sigma} s_1 + \cdots + \alpha_n \frac{1}{\lambda_n - \sigma} s_n$$

and

$$(A^{-1} - \sigma I)^k x = \alpha_1 \frac{1}{(\lambda_1 - \sigma)^k} s_1 + \cdots + \alpha_n \frac{1}{(\lambda_n - \sigma)^k} s_n$$

Exercise 11 *Show*

$$x_k = \alpha_1 \left(\frac{\lambda_n - \sigma}{\lambda_1 - \sigma}\right)^k s_1 + \alpha_2 \left(\frac{\lambda_n - \sigma}{\lambda_2 - \sigma}\right)^k s_n \cdots + \alpha_{n-1} \left(\frac{\lambda_n - \sigma}{\lambda_{n-1} - \sigma}\right)^k s_{n-1} + \alpha_n s_n$$

Exercise 12 *Show*

$$||x_k - \alpha_n s_n||_2 \leq (|\alpha_1| + \cdots + |\alpha_{n-1}|) \left(\frac{|\lambda_n - \sigma|}{|\lambda_{n-1} - \sigma|}\right)^k$$

and thus

$$||x_{k+1} - \alpha_n s_n||_2 \leq \left(\frac{|\lambda_n - \sigma|}{|\lambda_{n-1} - \sigma|}\right) (|\alpha_1| + \cdots + |\alpha_{n-1}|) \left(\frac{|\lambda_n - \sigma|}{|\lambda_{n-1} - \sigma|}\right)^k$$

Notice that thus the bound on the difference between x_k and $\alpha_n s_n$ (an eigenvector associated with λ_n) is decreased at each step by a factor $\frac{|\lambda_n - \sigma|}{|\lambda_{n-1} - \sigma|}$.

Notice that by picking $\sigma \approx \lambda_n$ this factor can be decreased essentially to $\frac{|\lambda_n - \sigma|}{|\lambda_{n-1} - \lambda_n|}$.

Note: one can dynamically pick σ by using $\sigma_i = \frac{x_i^\star A x_i}{x_i^\star x_i}$ (the Rayleigh quotient) at each step. This leads to the Rayleigh quotient iteration.

4 QR Algorithm

Consider the iteration

- $A_0 = A$
- do $i = 0, \ldots$ until convergence
  - $A_i \rightarrow Q_i R_i$
  - $A_{i+1} \leftarrow R_i Q_i$
- enddo

Exercise 13 *Show* $A_{i+1} = Q_i^\star A_i Q_i$ *and* $A_{i+1} = Q_i^\star Q_{i-1}^\star \cdots Q_0^\star A Q_0 \cdots Q_{i-1} Q_i$.

Exercise 14 *Show* $A^k = \hat{Q}_k \hat{R}_k$ *where* $\hat{Q}_k = Q_0 Q_1 \cdots Q_k$ *and* $\hat{R}_k = R_k R_{k-1} \cdots R_0$.

Exercise 15 *Let*

$$\hat{R}_k = \begin{pmatrix} \hat{\rho}_{11}^{(k)} & \hat{\rho}_{12}^{(k)} & \cdots & \hat{\rho}_{1n}^{(k)} \\ 0 & \hat{\rho}_{22}^{(k)} & \cdots & \hat{\rho}_{2n}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\rho}_{nn}^{(k)} \end{pmatrix}$$

and

$$\hat{Q}_k = \begin{pmatrix} \hat{q}_1^{(k)} & \hat{q}_2^{(k)} & \cdots & \hat{q}_n^{(k)} \end{pmatrix}$$

Show $A^k e_1 = \hat{\rho}_{11}^{(k)} \hat{q}_1^{(k)}$

Thus, the first column of $\hat{Q}_k$ is in the same direction as the vector that results from performing the power methods with $x = e_1$.

Exercise 16 *Show* $A^{-k} = \hat{R}_k^{-1} \hat{Q}_k^{\star}$.

Exercise 17 *Let*

$$\hat{R}_k^{-1} = \begin{pmatrix} \hat{\rho}_{11}^{(k)} & \tilde{\rho}_{12}^{(k)} & \cdots & \tilde{\rho}_{1n}^{(k)} \\ 0 & \tilde{\rho}_{22}^{(k)} & \cdots & \tilde{\rho}_{2n}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\rho}_{nn}^{(k)} \end{pmatrix}$$

Show $e_n^T A^{-k} = \tilde{\rho}_{11}^{(k)} \hat{q}_n^{(k)\,T}$

Conclude that the last column of $\hat{Q}_k$ converges to the eigenvector associated with the smallest eigenvalue of A.